

N 9 3 - 2 4 9 1 9

The Role of the World Data Centers in Handling Ocean Climate Data

Ferris Webster

Abstract

The World Data Center System, set up for the International Geophysical Year in 1957, is an international network of data centers that links data contributors to data users in the geosciences. It includes means for the synthesis, analysis, and preparation of data products. It was set up in response to the needs of the international scientific community, and is still overseen by non-governmental scientific organizations. Because it is freely available to researchers in all countries, the World Data Center System has a special role to play in support of ocean climate research and monitoring programs.

The World Data Centers face a number of challenges today. Apathy is probably the greatest, since many scientists take the system for granted. There is need to improve access and exploit new technology. The system must establish new links to assure continuity in a world with political changes. The multidisciplinary needs of global change research will demand capabilities for data and information management that go beyond the traditional emphasis on geophysics.

Introduction

Global ocean research programs and ocean observing systems need a robust and open system for collecting, processing, sharing, and archiving the abundant observations collected. The data and the products derived from them must be available to researchers in all countries.

Oceanographers, meteorologists, and other earth scientists are fortunate to have a system already in place for the international archiving, cataloguing, and exchange of data. It's the system of World Data Centers, established more than thirty years ago to serve the needs of the International Geophysical Year. It has continued ever since and today has a positive role to play in handling ocean climate data.

Strangely, there is nearly no mention of World Data Centers in the documents I have seen as background to this Ocean Climate Data Workshop. This is specially curious since this workshop is devoted to the subject of data. I am accustomed to finding data relegated to one of the final pages of any plan for global research. It's said that the emphasis should be on the science. I can't argue with that. However, much of the science will ultimately depend on the effective use of data. Thus I do

argue that an effective science plan will recognize the critical role of preserving and sharing data. In fact, the World Data Center system was established by scientists, to ensure that the vital task of preserving and sharing global data met the needs of scientists.

I would like to take the opportunity of this Workshop to describe what the World Data Center system is, provide a bit of its history, suggest what role it might play in ocean climate research and operational ocean programs, and point out what I believe are the greatest challenges the system faces today.

World Data Centers

"The World Data Center System is essentially a network of Data Centers linking data contributors to data users and includes means for the synthesis, analysis, and preparation of data products."—Guide to the World Data Center System

The World Data Centers (WDC) began as a network in 1957, to serve the needs of archiving and data exchange for the International Geophysical Year. The International Council of Scientific Unions (ICSU) took over the coordination of the system shortly thereafter. The network has functioned ever since, with oversight from ICSU's Panel on World Data Centers. The Centers serve the many disciplines of geophysical sciences and, in the oceanographic domain, include chemical and biological ocean data.

Note that ICSU is a non-governmental organization. The members of the ICSU Panel on World Data Centers are generally not individuals with responsibility for the operation of the system but are chosen from the scientific community.

The centers themselves are generally, but not exclusively, governmental. They are operated at the expense of the host countries, who have volunteered to maintain them. The WDC system thus has an interesting duality: the nationally operated and funded data centers are internationally coordinated by a nongovernmental body responding to the scientific community.

The United States operates a set of centers known as WDC-A, Russia operates WDC-B. China operates WDC-D. A number of centers in Europe, Japan, and India are known collectively as WDC-C. Most world data centers are collocated with related national data centers. In fact, at such centers there is generally considerable integration between the world and national data collections.

The WDCs operate under the terms of the Guide to the World Data Center System (1987). The Guide is reviewed every few years and issued by the ICSU Panel. Among some of the principles set forth in the guide, the world data centers should:

- operate for the benefit of the international scientific community.
- exchange data in all disciplines related to the Earth, its environment and the Sun.
- be completely accessible by scientists in all countries.
- exchange data among themselves.
- be supported by the host countries on a long-term basis.
- be coordinated within a country by the appropriate national committee or scientific institution.

In the USA, the world data centers and corresponding national data centers are:

<i>World Data Center A for</i>	<i>National</i>
Glaciology	Snow and Ice Data Center
Marine Geology and Geophysics	Geophysical Data Center
Meteorology	Climatic Data Center
Oceanography	Oceanographic Data Center
Rockets and Satellites	Space Science Data Center
Rotation of the Earth	
Seismology	Earthquake Information Center
Solar-Terrestrial Physics	Geophysical Data Center
Solid Earth Geophysics	Geophysical Data Center

In general, datasets are duplicated between data centers of the same type. Thus WDC-A, Oceanography, in Washington, DC, USA and WDC-B, Oceanography, in Obninsk, Kaluga, Russia, hold many duplicate datasets. Holding duplicates should assure preservation in case of disaster befalling one of the centers, and improve ease of access for users.

The datasets of most interest to the ocean climate community are held by the WDCs for Oceanography and Meteorology. However, datasets in the WDCs for Glaciology and Marine Geology and Geophysics as well as data available through the WDC for Rockets and Satellites may also be of interest for oceanographic needs.

World Data Centers and Ocean Climate Research and Monitoring

Ocean climate research and monitoring programs have been stimulated by concern over possible changes in the global environment. The programs are of global scale and measurements will be collected by investigators or by operational programs in many countries. The objective of understanding global change will likely only be achieved through collaboration between scientists, possibly of more than one discipline, and probably in more than one region of the world. Under such conditions, access to the datasets by researchers in all countries is of critical importance. Global measurements will be made in many ways, by many different kinds of researchers. If the scientific program is to succeed, an international

system to collect and distribute the data is necessary. The World Data Centers can meet this need.

The World Data Centers are responsible for collecting, distributing, and archiving geophysical datasets. The collecting activity assures that datasets will be aggregated so that maximum advantage can be taken of them. The distribution activity assures that researchers anywhere in the world can gain access to the data. The archiving activity can preserve a long-term environmental record.

International Exchange

"Data held by a WDC must be completely accessible by scientists in all countries, upon written request or personal visit." —ICSU Guide to the World Data Center System

For climate research programs, the World Data Centers present an valuable opportunity. The centers are already in place, they have experienced personnel, and they cover many of the disciplines relevant to global change research. The centers are international and have a tradition of working together and serving users in all countries. They provide a great base on which to build for the future.

The scope of the World Data Centers fits well with global climate research. Large-scale ocean climate research programs are global in scale. The participation of researchers in many countries is essential. Researchers need to find and share datasets with colleagues in other countries and the WDC system provides a means to meet that need.

The growth of ocean climate research programs in the coming decade will provide the WDCs with an opportunity to show their value. But we must be careful: we can't simply load more requirements on the WDCs and expect them to meet our needs. We should be discussing the role of the WDCs in meetings like this Workshop. There should be dialogue between the operators of the data centers and the planners of ocean climate research and monitoring programs. That dialogue should include a discussion of the role of the data system as well as a review of the resources that will be necessary to implement it.

In many countries (the United States is one) the world data centers generally have no specific budget allocation. They use facilities of the corresponding national center, created to serve national needs. The world data centers may have a limited role and a budget limited to what can be accomplished after national responsibilities are met. If we want to expand the role of the data centers to meet the needs of global-scale research, we ought to be sure that the resources are available.

I believe we should be asking the World Data Centers to meet the needs of international ocean climate research and monitoring programs. Let's put the

World Data Centers into the plans we are now creating. Let's use the World Data Centers to collect, archive, and distribute data and data products for researchers in all countries.

I also believe we should choose World Data Centers over bilateral data-sharing arrangements. There is an increasing use of bilateral agreements for data-sharing between countries, particularly for satellite data. Though bilateral arrangements may be easier to establish, they may not be best for global research. They are generally closed to "outsiders". They are generally set up by governments, and are subject to national restrictions. For true international progress in understanding global climate, we need a system that is international, open to all, and set up to meet the needs of research. The existing World Data Center System fits the bill.

Challenges Facing the WDC System

A number of challenges face the World Data Center System three decades after its creation. The system has worked well during this time. However, unless it continues to evolve to meet changing scientific needs it will likely stagnate and ultimately become irrelevant.

Apathy

The number one problem is apathy. Many environmental scientists take the system for granted or are unaware of its existence. Scientific involvement is essential. The WDC system will preserve its unique advantages only if the scientific community continues to support it.

If I can send one message to the researchers working on ocean climate programs it is this: Get involved, define your needs for datasets, and work to strengthen the World Data Center System to meet those needs.

Responding to Political Changes

Political changes in the former Soviet Union are having an impact on the current operations and future outlook for WDC-B. Since it is a key link in the system, WDC-B should be preserved and strengthened. One of the inspiring triumphs of the WDC system was that the principle of free access by all scientists had been respected by both sides throughout the cold war. Current political changes and realignments in the former Soviet Union could jeopardize operations and even the continuity of the data holdings of WDC-B. The international system will be challenged to protect and continue WDC-B.

Multidisciplinary Needs of Global Change Research

The current WDC system focuses on the igy disciplines. The geosciences have evolved considerably, but only a few new types of world data centers have been

created since the original network was set up. Global change research objectives involve understanding interactions between processes that are traditionally treated by differing disciplines. Thus, to respond to global change research, databases should be incorporated into the system in disciplines so far not covered. Some examples are in terrestrial and oceanic biology, atmospheric chemistry, hydrology, and land processes. As a result of current discussions, I am hopeful that many of these "new" disciplines will be served by world data centers to be set up in the next few years.

Strengthen Metadata Standards

Many important historical datasets are limited by the information we have about them: when they were collected, where they were collected, what instrument was used to make the measurement, what the calibration of the instrument was, etc.

Such information about data is often referred to as "metadata". Without the proper metadata, many datasets are not usable. Nevertheless, nationally and internationally, insufficient emphasis is placed on providing the metadata along with the data. Data centers have too often in the past been negligent in ensuring that the metadata is linked inextricably with the corresponding data.

As we develop stronger long-term climate research programs, the critical importance of metadata will become even more evident. The data centers of the WDC system will be challenged to strengthen the standards for metadata and to ensure their incorporation into standard operations.

Data Rescue

Long-term time series are critical for climate research. A good data point collected fifty years ago could be just as important to establishing climate trends as a data point collected fifty years from now. Consequently, we should identify existing datasets, and assure that they and the corresponding metadata are preserved.

Finding and rescuing datasets is a major challenge. Worldwide, many environmental datasets are in poor physical condition. Datasets are dispersed and many valuable measurements are "lost" and in danger of being destroyed through ignorance of their value. Some data were collected by one-of-a-kind instruments. Others have been processed by computer systems that are no longer in existence. As a consequence, these datasets are often in formats that are practically inaccessible.

The world data centers have the opportunity to play a key role in identifying and rehabilitating historical datasets. A major international program of data rescue should be sponsored by the WDC system.

Products and Information

Many users of the WDC system are not simply interested in raw datasets. Rather, they prefer data products or services derived from the data. With the growing interest in the environment, there is an increasing demand for information rather than data. Here again, products or analyses based on datasets are the preferred item. To satisfy this demand, the system should strive to develop new products that respond to the needs of the global change research community.

To go along with this approach, the data centers will have to develop information management capabilities. This change would broaden the role of the data centers and transform them into data and information centers. Incorporating the information analysis role would respond to the evolving needs of the research community and maintain the relevance of the data centers.

New Technology

The system must evolve to incorporate new technology. New means of data exchange and communication are transforming the ability of researchers to find and obtain datasets. New kinds of datasets are emerging, especially from remote sensing, which require new technologies for processing, storage, cataloguing, and distribution. Many of the individual world data centers have taken the lead in showing how these new technologies can be advantageously exploited. Their techniques should be made available to all. In an international network, it is desirable to have the capabilities incorporated into all centers.

Access to the WDC archives should be improved. Letters and visits are no longer the only way for scientists to communicate with data centers. For example, international computer network access is essential. On-line catalog systems should be developed. Data dissemination by diskette, by CD ROMs and by other new technologies should be made standard options.

Ocean monitoring and prediction systems will add a new challenge to the operations of the oceanographic data centers. Oceanographic data centers have traditionally received a relatively small amount of their data from operational systems. This is in contrast to the meteorological data centers, where the bulk of the data is operational. Serving ocean monitoring and prediction programs will call for developing new procedures for data handling and additional ocean data products.

Conclusion

The purpose of this Workshop is to lead to the improved data delivery systems needed by researchers studying the ocean's role in climate change. To meet that end, the ocean science community should incorporate the World Data Centers

into its plans. Along with that, there should be a renewed commitment to the support and improvement of the World Data Centers.

Appendix: The World Data Center Network

WDC-A

Oceanography	Washington
Meteorology	Asheville
Glaciology	Boulder
Marine geology and geophysics	Boulder
Rockets and satellites	Greenbelt
Rotation of the Earth	Washington
Seismology	Golden
Solar-terrestrial physics	Boulder
Solid-earth geophysics	Boulder

WDC-B1

Oceanography	Obninsk
Meteorology	
Marine geology and geophysics	
Glaciology	
Rockets and satellites	
Rotation of the Earth	
Tsunami, mean sea level, tides	

WDC-B2

Solar-terrestrial physics	Moscow
Solid-earth geophysics	

WDC-C1

Earth tides	Brussels
Geomagnetism	Copenhagen
Geomagnetism	Edinburgh
Recent crustal movements	Prague
Solar Activity	Meudon
Solar-terrestrial physics	Chilton
Sunspot index	Brussels
Glaciology	Cambridge
Soil geography & classification	Wageningen

WDC-C2

Airglow	Tokyo
Aurora	Kaga
Cosmic Rays	Nagoya
Geomagnetism	Kyoto
Geomagnetism	Bombay
Ionosphere	Tokyo

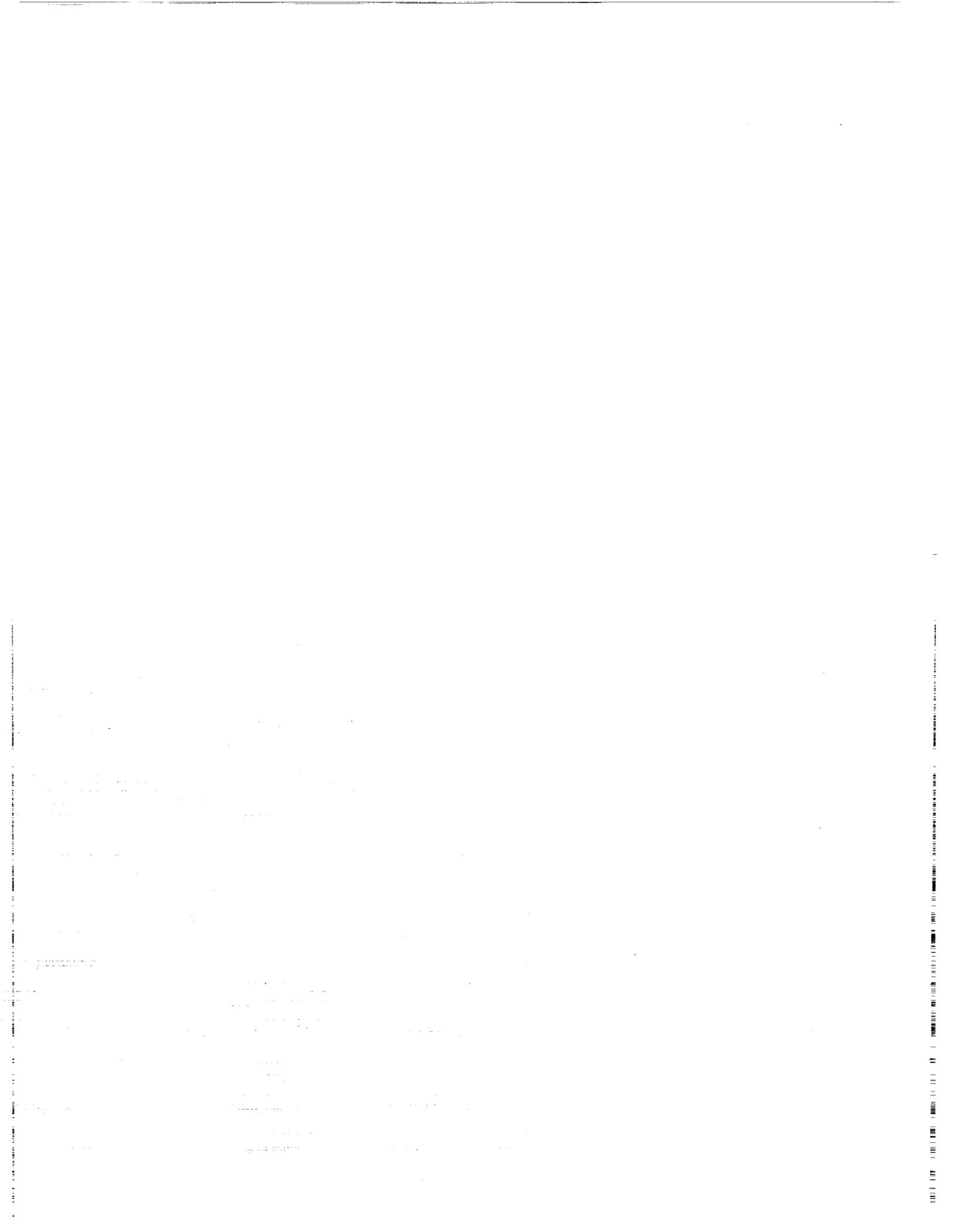
Nuclear radiation
Solar radio emissions
Solar-terrestrial activity

Tokyo
Toyokawa
Sagamihara

WDC-D

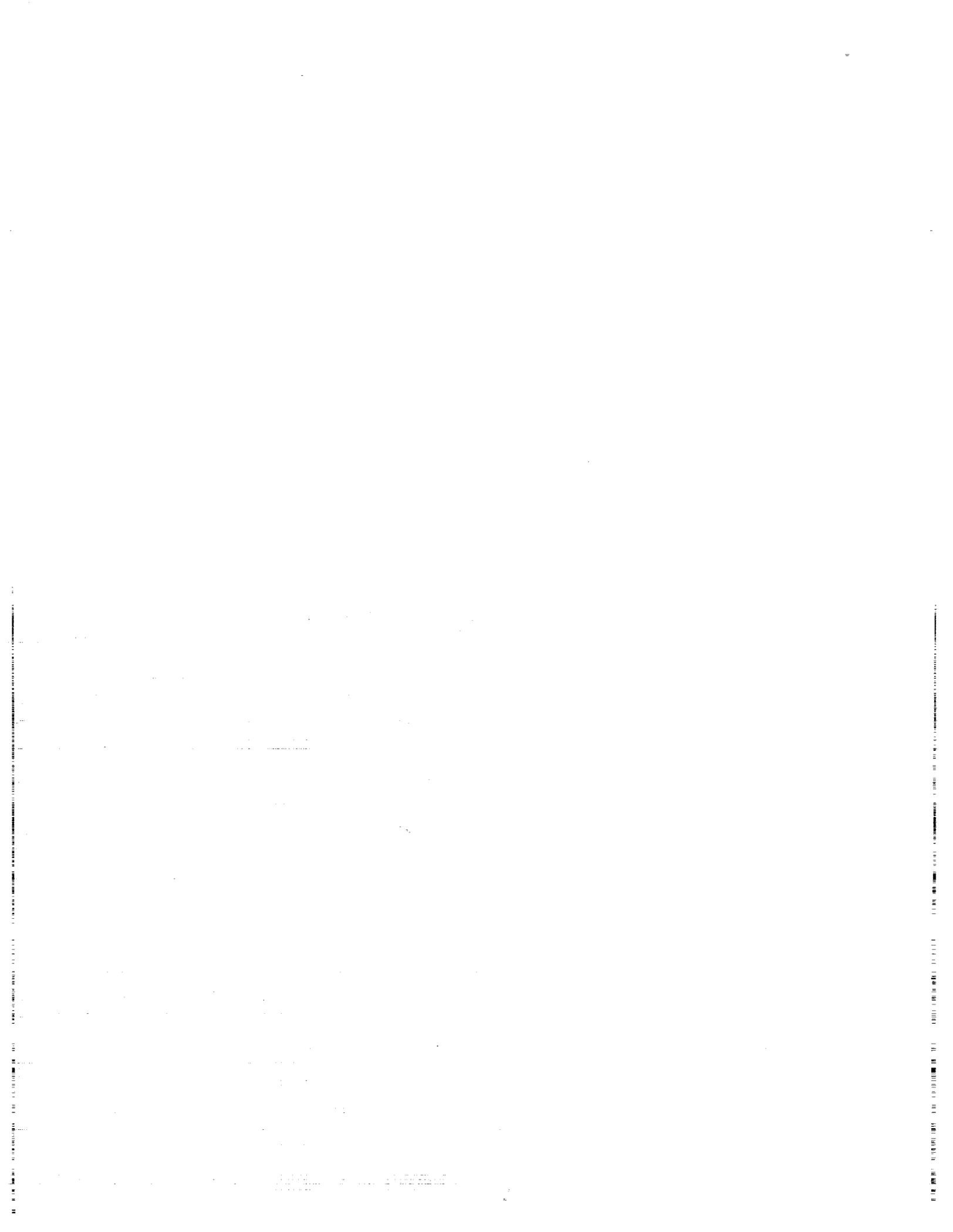
Oceanography
Meteorology
Seismology
Geology
Renewable resources & environment
Astronomy
Glaciology & geocryology
Geophysics

Tianjin
Beijing
Beijing
Beijing
Beijing
Beijing
Lanzhou
Beijing



Monitoring Changes in the Ocean and Atmosphere

Convener - J. Ronald Wilson



Monitoring Changes in the Ocean and Atmosphere

The papers presented in this session provided insight into experiences gained in both field experiments and in modeling in recent years. Both in situ and remotely sensed data were discussed in some detail. The paper on the Global Ocean Observing System (GOOS) gave participants an opportunity to see how some of what we have learned is being applied to GOOS planning and to suggest ways in which these could be improved. The IOC sponsored Ocean PC project added the dimension of distributing data and data products to a large number of countries, especially the developing countries. Dr. Fiuza was able to present the view of an ultimate user of monitoring products and models.

Although the session was primarily aimed at rapid delivery of data from many sources to modelers and to users of contemporaneous data, a number of questions were raised regarding the archiving of data used in monitoring the oceans and resulting model output. Discussions and questions covered a host of topics including data acquisition, communications, quality control, data search and retrieval, formats used and needed, data products and data delivery. A number of the issues and recommendations listed in the "Wrap-up" section of these proceedings were derived from these papers and ensuing discussions.

Although the full text of Dr. Leetmaa's paper is not available at time of publication, he was able to provide us with a summary of his remarks.

